# Local Level Model

Siem Jan Koopman

http://sjkoopman.net
s.j.koopman@vu.nl
Department of Econometrics, Vrije Universiteit Amsterdam
CREATES, Aarhus University

October 2016: CREATES, Aarhus University

# Day 1 : Local Level Model

Program :

- Introduction
- Local level model
- Statistical dynamic properties
- Filtering, smoothing and forecasting.
- Literature : J. Durbin and S.J. Koopman (2012), "Time Series Analysis by State Space Methods", Second Edition, Oxford: Oxford University Press. Chapter 2.
- Exercises and Assignments.

# Time Series

A time series is a set of observations $y_t$, each one recorded at a specific time $t$.

The observations are ordered over time.
We assume to have $n$ observations, $t = 1, \ldots, n$.

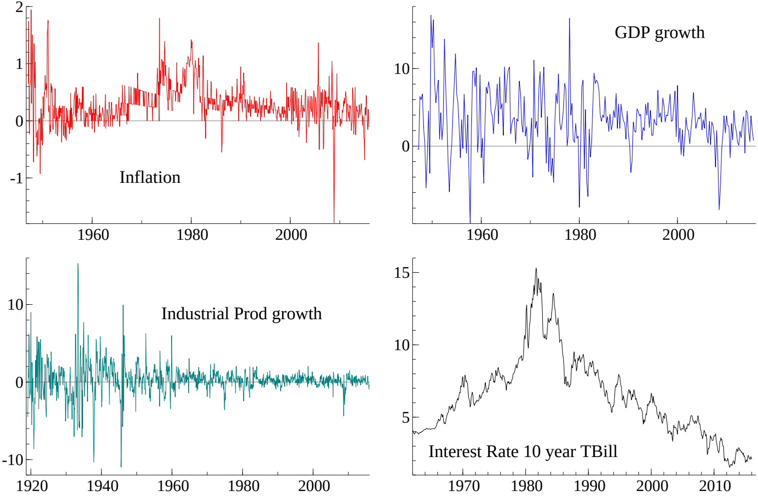
Examples of time series are:

- Number of cars sold each year
- Gross Domestic Product of a country
- Stock prices during one day
- Number of firm defaults

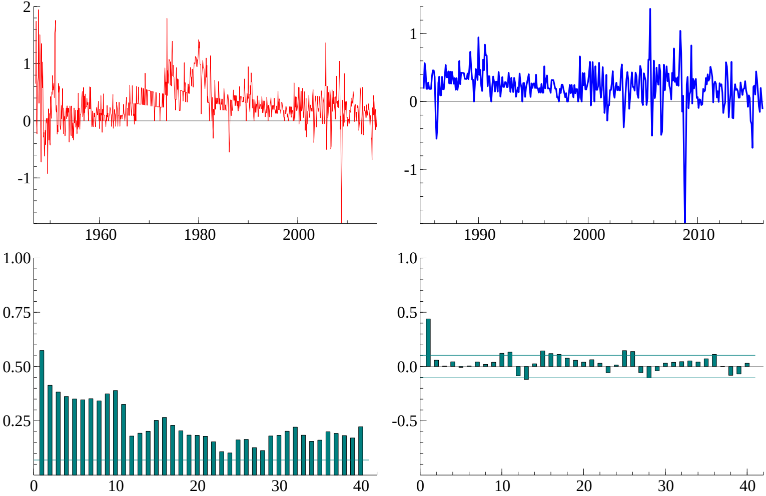
Our purpose is to identify and to model the serial or "dynamic" correlation structure in the time series.

Time series analysis is relevant for a wide variety of tasks including economic policy, financial decision making and forecasting

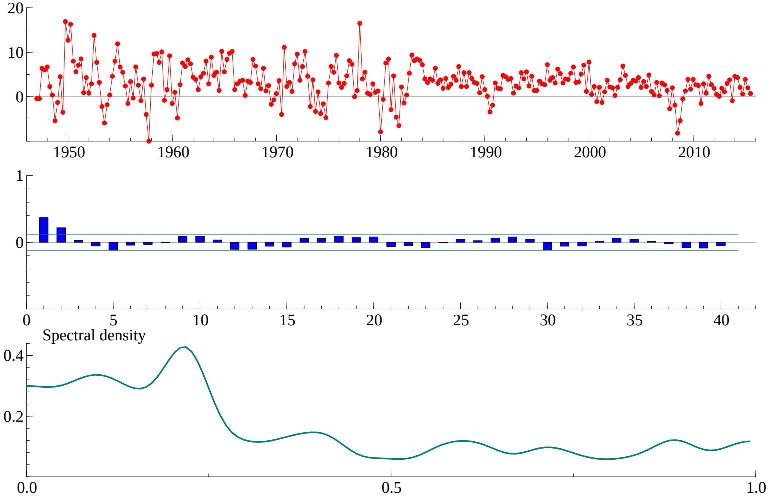
# The US Economy



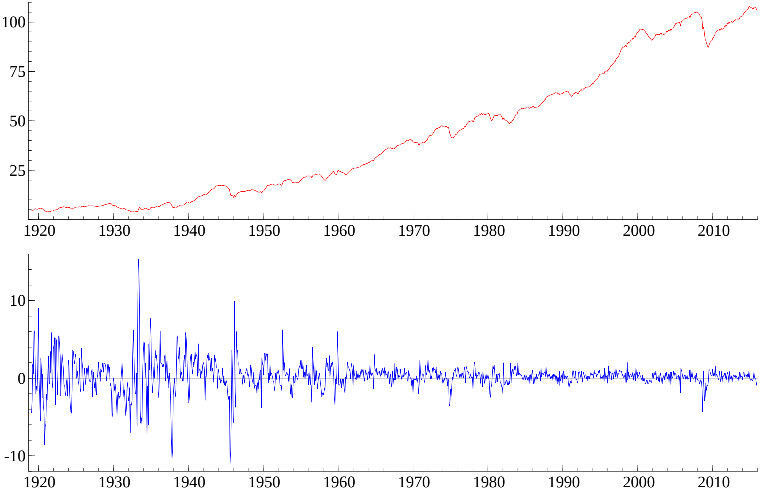
# US Inflation, based on CPI, all products



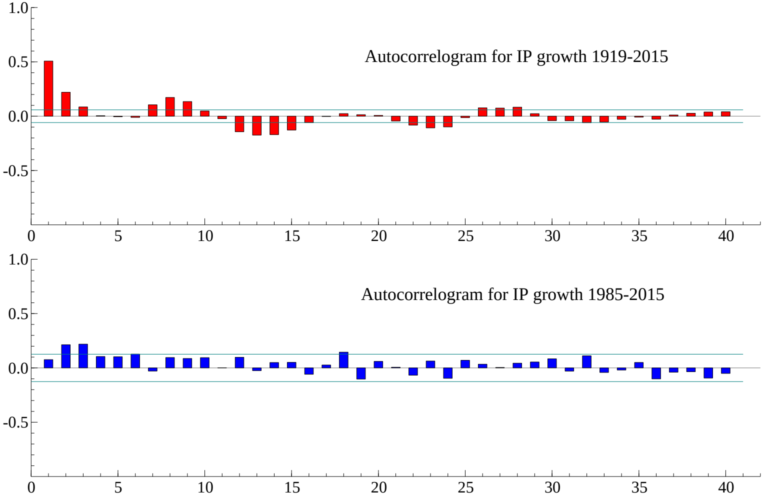
# US Gross Domestic Product (GDP), percentage growth



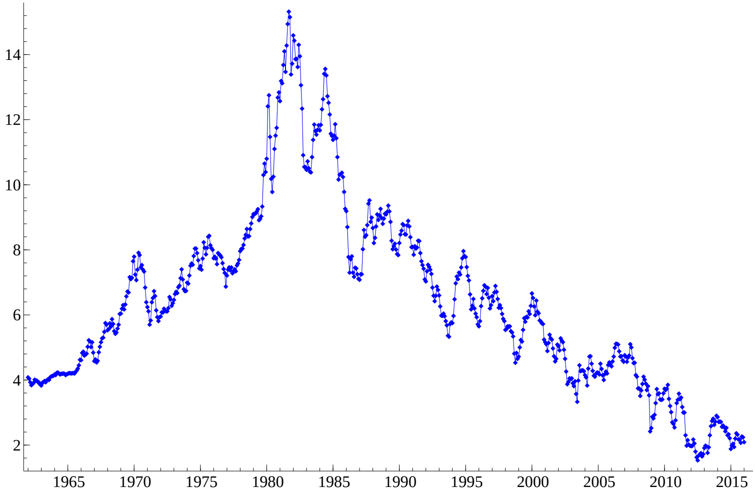
# US Industrial Production, levels and growth



# US Industrial Production, growth



# US Treasury Bill Rate, 10 years

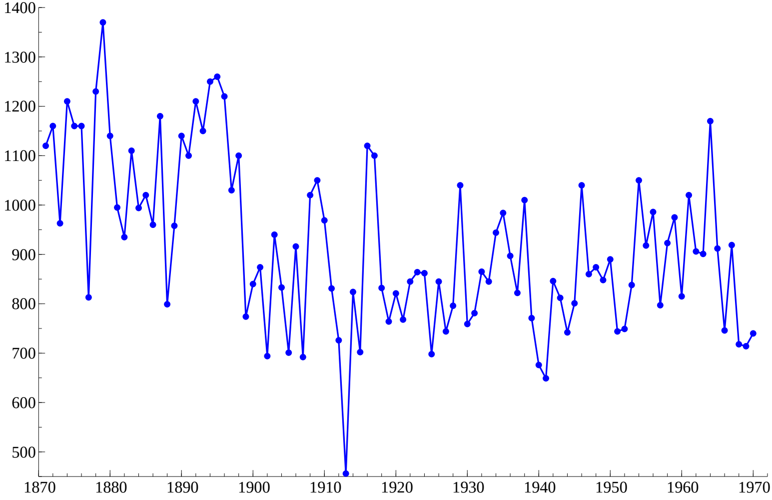


# Sources of time series data

Data sources :

- US economics :
  http://research.stlouisfed.org/fred2/
- DK book data :
  http://www.ssfpack.com/files/DK-data.zip
- Financial data : Datastream, Wharton, Yahoo Finance
- Time Series Data Library of Rob Hyndman :
  http://datamarket.com/data/list/?q=provider:tsdl

# Example: Nile data, yearly volumes



# Time Series

A time series for a single entity is typical denoted by

$$y_1, \ldots, y_n \quad \Leftrightarrow \quad y_t, \qquad t = 1, \ldots, n,$$

where $t$ is the time index and $n$ is time series length.
The current value is $y_t$.
The first lagged value, or first lag, is $y_{t-1}$.
The $\tau$th lagged value, or $\tau$-th lag, is $y_{t-\tau}$ for $\tau = 1, 2, 3, \ldots$.

The change between period $t-1$ and period $t$ is $y_t - y_{t-1}$.
This is called the first difference denoted by $\Delta y_t = y_t - y_{t-1}$.

In economic time series, we often take the first difference of the logarithm, or the log-difference, that is

$$\Delta \log y_t = \log y_t - \log y_{t-1} = \log(y_t / y_{t-1}),$$

is a proxy of proportional change, see Appendix.
Percentage change is then $100\Delta \log y_t$.

# Time Series Models: many

- Autoregressive models, unit roots
- Autoregressive moving average models
- Long memory models, fractional integration
- ... *unobserved components time series models* ...
- Dynamic regression models, error correction models
- Vector autoregressive models, cointegration, vector error correction models
- ... *state space models* ...
- Regime-switching, Markov-switching, treshold autoregression, smooth transitions models
- Generalized autoregressive conditional heteroskedasticity (GARCH) models
- Autoregressive conditional duration models and related models
- ... *stochastic volatility models* ...

# Autoregressive model: AR(1)

The AR(1) model is given by

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2),$$

with three parameter coefficients $\mu$, $\phi$ and $\sigma_\varepsilon^2$ with $0 < \sigma_\varepsilon < \infty$.

Stationary condition: $|\phi| < 1$.

Statistical dynamic properties:

- Mean $\mathbb{E}(y_t) = \mu \,/\, (1 - \phi)$; in case $\mu = 0$, $\mathbb{E}(y_t) = 0$;
- Variance $\mathbb{V}\text{ar}(y_t) = \sigma^2 \,/\, (1 - \phi^2)$;
- Autocovariance lag 1 is $\mathbb{C}\text{ov}(y_t, y_{t-1}) = \phi\, \sigma^2 \,/\, (1 - \phi^2)$;
- and for lag $\tau = 2, 3, 4, \ldots$ is $\mathbb{C}\text{ov}(y_t, y_{t-\tau}) = \phi^\tau \sigma^2 \,/\, (1 - \phi^2)$;
- Autocorrelation lag $\tau = 1, 2, 3, \ldots$ is $\mathbb{C}\text{orr}(y_t, y_{t-\tau}) = \phi^\tau$.

# Moving Average model: MA(1)

The MA(1) model is given by

$$y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2),$$

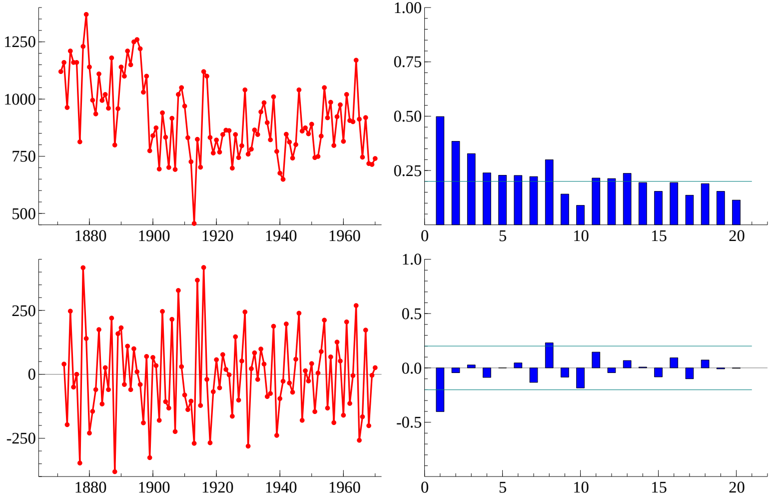
with three parameter coefficients $\mu$, $\theta$ and $\sigma_\varepsilon^2$ with $0 < \sigma_\varepsilon < \infty$.

Invertibility condition: $|\theta| < 1$.

Statistical dynamic properties:

- Mean $\mathbb{E}(y_t) = \mu$; in case $\mu = 0$, $\mathbb{E}(y_t) = 0$;
- Variance $\mathbb{V}\text{ar}(y_t) = \sigma^2\,(1 + \theta^2)$;
- Autocovariance lag 1 is $\mathbb{C}\text{ov}(y_t, y_{t-1}) = \theta\,\sigma^2$;
- ... for lag $\tau = 2, 3, 4, \ldots$ is $\mathbb{C}\text{ov}(y_t, y_{t-\tau}) = 0$;
- Autocorrelation lag 1 is $\mathbb{C}\text{orr}(y_t, y_{t-1}) = \theta\,/\,(1 + \theta^2)$.

# Example: Nile in levels and Nile in differences



# Classical Decomposition

A basic model for representing a time series is the additive model

$$y_t = \mu_t + \gamma_t + \psi_t + \varepsilon_t, \qquad t = 1, \ldots, n,$$

also known as the Classical Decomposition:

$$\begin{aligned} y_t &= \text{observation}, \\ \mu_t &= \text{slowly changing component (trend)}, \\ \gamma_t &= \text{periodic component (seasonal)}, \\ \psi_t &= \text{stationary component (cycle, ARMA)}, \\ \varepsilon_t &= \text{irregular component (disturbance)}. \end{aligned}$$

It is an *Unobserved Components* time series model, when the components are modelled as dynamic stochastic processes.

# Local Level Model

- Component is stochastic or deterministic function of time:
  - Deterministic, eg: $y_t = \mu(t) + \varepsilon_t$ with $\varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2)$
  - Stochastic, eg: *Local Level* model:
- Local level model :

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, \qquad & \varepsilon_t &\sim \mathcal{NID}(0, \sigma_\varepsilon^2) \\ \mu_{t+1} &= \mu_t + \eta_t, \qquad & \eta_t &\sim \mathcal{NID}(0, \sigma_\eta^2) \end{aligned}$$

  - The disturbances $\varepsilon_t$, $\eta_s$ are independent for all $s, t$;
  - The model is incomplete without initial specification for $\mu_1$.
  - The time series processes for $\mu_t$ and $y_t$ are nonstationary.

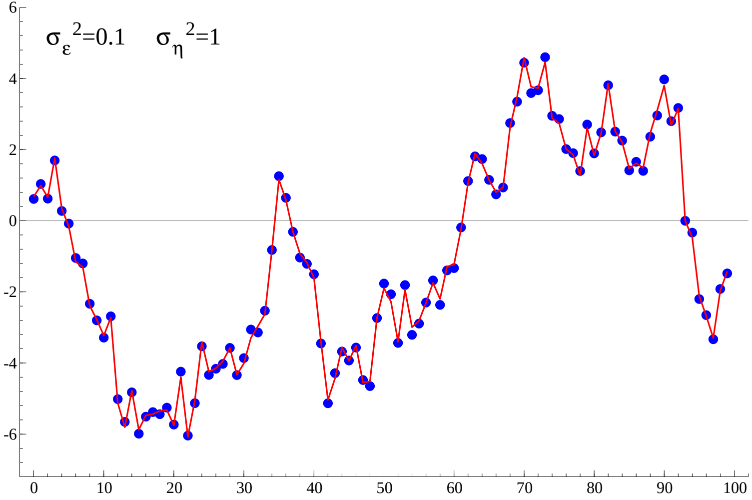
# Local Level Model

The local level model or random walk plus noise model :

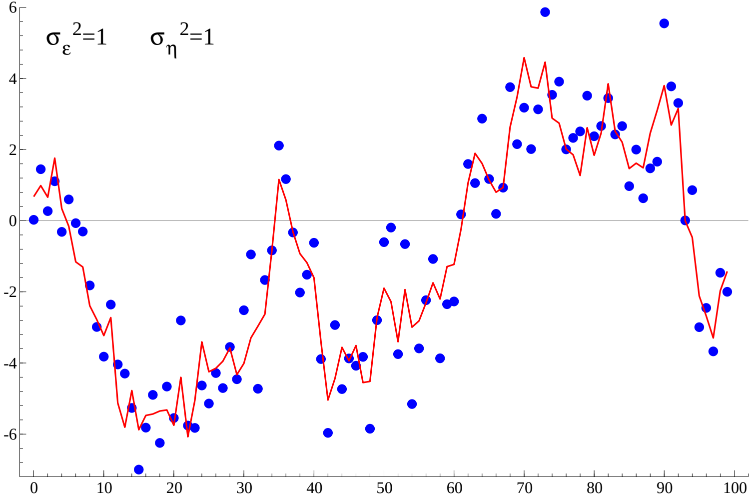
$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2)$$
$$\mu_{t+1} = \mu_t + \eta_t, \qquad \eta_t \sim \mathcal{NID}(0, \sigma_\eta^2)$$

- The level $\mu_t$ and irregular $\varepsilon_t$ are unobserved;
- Parameters $\sigma_\varepsilon^2$ and $\sigma_\eta^2$ are unknown;
- We still need to define $\mu_1$;
- Trivial special cases:
  - $\sigma_\eta^2 = 0 \implies y_t \sim \mathcal{NID}(\mu_1, \sigma_\varepsilon^2)$ (IID constant level);
  - $\sigma_\varepsilon^2 = 0 \implies y_{t+1} = y_t + \eta_t$ (random walk);
- Local Level model is basic illustration of state space model.

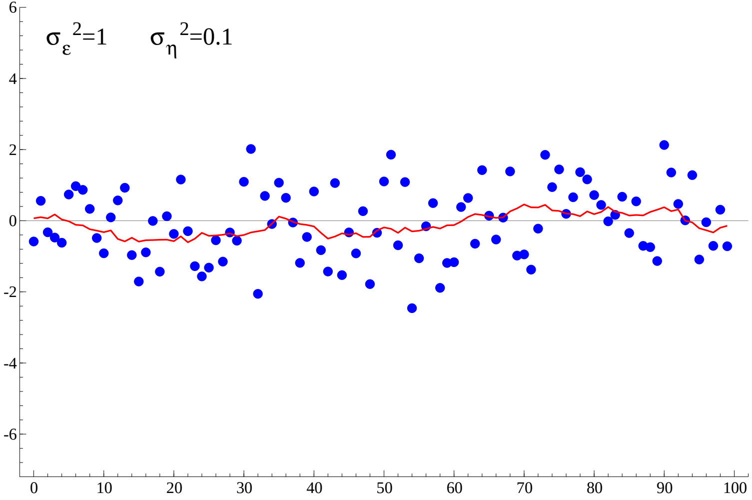
# Simulated Local Level data



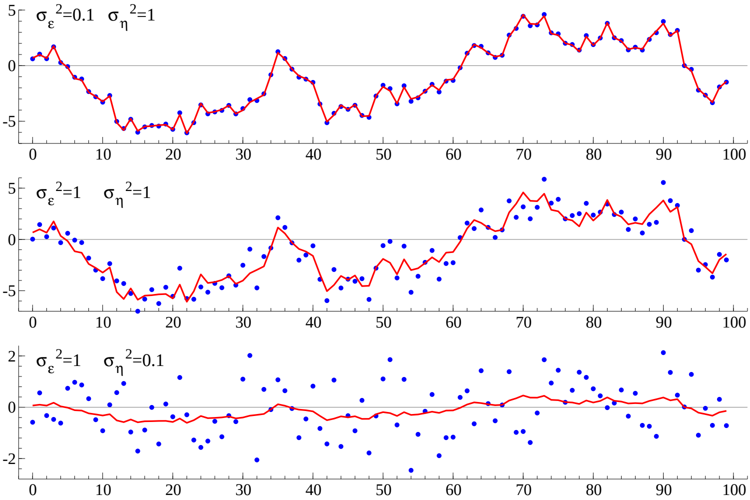
# Simulated Local Level data



# Simulated Local Level data



# Simulated Local Level data



# Properties of Local Level model

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2),$$
$$\mu_{t+1} = \mu_t + \eta_t, \qquad \eta_t \sim \mathcal{NID}(0, \sigma_\eta^2),$$

- First difference is stationary:

$$\Delta y_t = \Delta \mu_t + \Delta \varepsilon_t = \eta_{t-1} + \varepsilon_t - \varepsilon_{t-1}.$$

- Dynamic properties of $\Delta y_t$:

$$\begin{aligned}
\mathbb{E}(\Delta y_t) = 0,& \\
\gamma_0 &= \mathbb{E}(\Delta y_t \Delta y_t) = \sigma_\eta^2 + 2\sigma_\varepsilon^2, \\
\gamma_1 &= \mathbb{E}(\Delta y_t \Delta y_{t-1}) = -\sigma_\varepsilon^2, \\
\gamma_\tau &= \mathbb{E}(\Delta y_t \Delta y_{t-\tau}) = 0 \quad \text{for } \tau \geq 2.
\end{aligned}$$

## Properties of Local Level model

- Define $q$ as the *signal-to-noise ratio* : $q = \sigma^2_\eta \,/\, \sigma^2_\varepsilon$
- The theoretical ACF of $\Delta y_t$ is

$$\rho_1 = \frac{-\sigma^2_\varepsilon}{\sigma^2_\eta + 2\sigma^2_\varepsilon} = -\frac{1}{q+2},$$

$$\rho_\tau = 0, \qquad \tau \geq 2.$$

- It implies that

$$-1/2 \leq \rho_1 \leq 0$$

- The local level model implies that $\Delta y_t \sim$ MA(1). Hence $y_t \sim$ is ARIMA$(0,1,1)$. We have $\Delta y_t = \xi_t + \theta \xi_{t-1}, \quad \xi_t \sim \mathcal{NID}(0, \sigma^2)$.
- This implied MA(1) has ACF $\rho_1 = \theta \,/\, (1 + \theta^2)$, and hence a restricted parameter space for $\theta$ : $-1 < \theta < 0$.
- To express $\theta$ as function of $q$, solve equality for $\rho_1$'s:

$$\theta = \frac{1}{2}\left(\sqrt{q^2 + 4q} - 2 - q\right).$$

# Local Level Model

The Local Level model is given by

$$y_t = \mu_t + \varepsilon_t, \qquad \mu_{t+1} = \mu_t + \eta_t, \qquad t = 1, \ldots, n.$$

- The parameters $\sigma^2_\varepsilon$ and $\sigma^2_\eta$ are unknown and need to be estimated, typically via maximum likelihood estimation;
- MLE for this class of models is discussed in next session.
- When we treat parameters $\sigma^2_\varepsilon$ and $\sigma^2_\eta$ as known, how to "estimate" the unobserved series $\mu_1, \ldots, \mu_n$ ?
- This "estimation" is referred to as signal extraction.
- We base this "estimation" on conditional expectations.
- Signal extraction is the recursive evaluation of conditional means and variances of the unobserved $\mu_t$ for $t = 1, \ldots, n$.
- It is known as the Kalman filter;
- Next we provide the derivation only for the Local Level model.
- In our next session we discuss the Kalman filter for the general linear state space model.

# Signal extraction: conditional expectation

Consider two random variable $x$ and $y$ that are normally distributed

$$x \sim \mathcal{N}(\mu_x, \sigma_x^2), \qquad y \sim \mathcal{N}(\mu_y, \sigma_y^2), \qquad \mathbb{C}\text{ov}(x, y) = \sigma_{xy}.$$

Assume that we do not know anything about $x$ but we have collected an observation for $y$.

The conditional expectation and variance are given by

$$\mathbb{E}(x|y) = \mu_x + \sigma_{xy}(y - \mu_y) \,/\, \sigma_y^2, \qquad \mathbb{V}\text{ar}(x|y) = \sigma_x^2 - \sigma_{xy}^2 \,/\, \sigma_y^2.$$

Verify these results and make sure you can derive these results from basic principles. We have

$$x|y \sim f(x|y) \equiv \mathcal{N}(\mu_{x|y}, \sigma_{x|y}^2),$$

where $\mu_{x|y} \equiv \mathbb{E}(x|y)$ and $\sigma_{x|y}^2 \equiv \mathbb{V}\text{ar}(x|y)$.

Notice that $\mu_{x|y}$ is a function of $y$ but $\sigma_{x|y}^2$ is not.

Notice that when $\sigma_{xy} = 0$, $\mathbb{E}(x|y) = \mu_x$ and $\mathbb{V}\text{ar}(x|y) = \sigma_x^2$.

# Local Level Model: signal extraction

Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

Assume we have collected observations for $y_1, \ldots, y_{t-1}$ and that the conditional density $f(\mu_t | y_1, \ldots, y_{t-1})$ is normal with known mean $a_t$ and known variance $p_t$, we have

$$\mu_t | y_1, \ldots, y_{t-1} \sim f(\mu_t | y_1, \ldots, y_{t-1}) \equiv \mathcal{N}(a_t, p_t).$$

Next we collect an observation for $y_t$, the conditional densities of interest are

$$f(\mu_t | y_1, \ldots, y_t), \qquad f(\mu_{t+1} | y_1, \ldots, y_t).$$

These conditional densities turn out to be normal as well

$$f(\mu_t | y_1, \ldots, y_t) \equiv \mathcal{N}(a_{t|t}, p_{t|t}), \quad f(\mu_{t+1} | y_1, \ldots, y_t) \equiv \mathcal{N}(a_{t+1}, p_{t+1}).$$

Can we express $(a_{t|t}, p_{t|t})$ in terms of $(a_t, p_t)$ ? Also $(a_{t+1}, p_{t+1})$ ?

# Local Level Model: signal extraction

Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2_\varepsilon), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma^2_\eta).$$

Notation: $Y_s = \{y_1, \ldots, y_s\}$, for $s = t-1$, $s = t$ and $s = n$.

Define prediction error $v_t = y_t - a_t$ with $a_t = \mathbb{E}(\mu_t | Y_{t-1})$, with properties such as

$$\begin{aligned}
\mathbb{E}(v_t | Y_{t-1}) = \mathbb{E}(\mu_t + \varepsilon_t - a_t | Y_{t-1}) &= a_t - a_t = 0, \\
\mathbb{V}\mathrm{ar}(v_t | Y_{t-1}) = \mathbb{V}\mathrm{ar}(\mu_t - a_t + \varepsilon_t | Y_{t-1}) &= p_t + \sigma^2_\varepsilon, \\
\mathbb{E}(v_t | \mu_t, Y_{t-1}) &= \mu_t - a_t, \\
\mathbb{V}\mathrm{ar}(v_t | \mu_t, Y_{t-1}) &= \sigma^2_\varepsilon,
\end{aligned}$$

We have $\mathbb{E}(\varepsilon_t) = 0$ but verify that $\mathbb{E}(\varepsilon_t | Y_{t-1}) = 0$.
When $y_t$ is observed, it becomes fixed, just as $y_1, \ldots, y_{t-1}$.
But also $v_t$ is then fixed, it is non-stochastic !!

# Local Level Model: signal extraction

Next, we aim to obtain an expression for $f(\mu_t | y_1, \ldots, y_t)$, with an eye on updating. Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

Consider filtered estimate $f(\mu_t | y_1, \ldots, y_t) \equiv f(\mu_t | v_t, Y_{t-1})$ since $v_t = y_t - a_t$, where $a_t = \mathbb{E}(\mu_t | Y_{t-1})$, are all fixed. We have

$$\begin{aligned} f(\mu_t | v_t, Y_{t-1}) &= f(\mu_t, v_t | Y_{t-1}) / f(v_t | Y_{t-1}) \\ &= f(\mu_t | Y_{t-1}) f(v_t | \mu_t, Y_{t-1}) / f(v_t | Y_{t-1}), \end{aligned}$$

where $f()$'s are normals and $f(\mu_t | Y_t) = \text{const.} \times \exp\left(-\frac{1}{2} Q_t\right)$ with

$$Q_t = (\mu_t - a_t)^2 / p_t + (v_t - \mu_t + a_t)^2 / \sigma_\varepsilon^2 - v_t^2 / (p_t + \sigma_\varepsilon^2).$$

After some algebra, we have

$$Q_t = \frac{p_t + \sigma_\varepsilon^2}{p_t \, \sigma_\varepsilon^2} \left(\mu_t - a_t - \frac{p_t \, v_t}{p_t + \sigma_\varepsilon^2}\right)^2.$$

# Local Level Model: signal extraction

Next we consolidate these results for the Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

We are interested in the *filtered* signal density

$$f(\mu_t | Y_t) = \text{const.} \times \exp\big( -\frac{1}{2} Q_t \big),$$

with

$$Q_t = \frac{p_t + \sigma_\varepsilon^2}{p_t \, \sigma_\varepsilon^2} \big( \mu_t - a_t - \frac{p_t \, v_t}{p_t + \sigma_\varepsilon^2} \big)^2.$$

It implies that

$$f(\mu_t | Y_t) \equiv \mathcal{N}(a_{t|t}, p_{t|t}),$$

with

$$a_{t|t} = a_t + k_t \, v_t, \qquad p_{t|t} = k_t \, \sigma_\varepsilon^2, \qquad k_t = \frac{p_t}{p_t + \sigma_\varepsilon^2}.$$

# Local Level Model: signal extraction

Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

In addition, we are typically interested in the *predicted* signal density

$$f(\mu_{t+1}|Y_t) \equiv \mathcal{N}(a_{t+1}, p_{t+1}),$$

where

$$a_{t+1} = \mathbb{E}(\mu_{t+1}|Y_t) = \mathbb{E}(\mu_t + \eta_t|Y_t) = a_{t|t},$$
$$p_{t+1} = \mathbb{V}\text{ar}(\mu_t + \eta_t|Y_t) = p_{t|t} + \sigma_\eta^2.$$

We have obtained the updating equations

$$a_{t+1} = a_t + k_t\, v_t, \qquad p_{t+1} = k_t\, \sigma_\varepsilon^2 + \sigma_\eta^2, \qquad k_t = \frac{p_t}{p_t + \sigma_\varepsilon^2}.$$

This is the celebrated Kalman filter for the Local Level model.

# Kalman filter for the Local Level Model

Local Level model :

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

The Kalman filter equations are given by

$$\begin{aligned}
v_t &= y_t - a_t, \qquad \mathbb{V}\mathrm{ar}(v_t) = p_t + \sigma_\varepsilon^2, \\
k_t &= p_t \,/\, (p_t + \sigma_\varepsilon^2), \\
a_{t|t} &= a_t + k_t v_t, \\
p_{t|t} &= k_t\, \sigma_\varepsilon^2, \\
a_{t+1} &= a_{t|t}, \\
p_{t+1} &= p_{t|t} + \sigma_\eta^2,
\end{aligned}$$

for $t = 1, \ldots, T$ with initialisation ...

# Kalman filter for the Local Level Model

Local Level model :

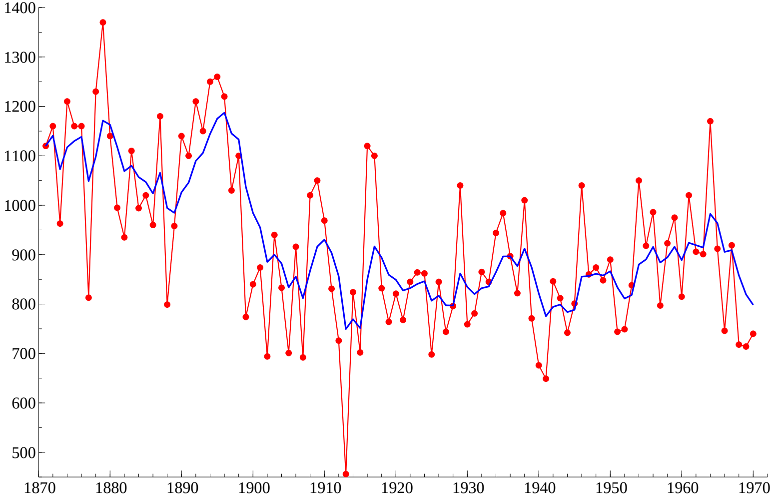
$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2).$$

The Kalman filter equations are given by

$$\begin{aligned}
v_t &= y_t - a_t, \qquad \mathbb{V}\text{ar}(v_t) = p_t + \sigma_\varepsilon^2, \\
k_t &= p_t \,/\, (p_t + \sigma_\varepsilon^2), \\
a_{t|t} &= a_t + k_t v_t, \\
p_{t|t} &= k_t \, \sigma_\varepsilon^2, \\
a_{t+1} &= a_{t|t}, \\
p_{t+1} &= p_{t|t} + \sigma_\eta^2,
\end{aligned}$$

for $t = 1, \ldots, T$ with initialisation ... $a_1 = 0$ and $p_1 = \sigma_\varepsilon^2 \times 10^7$.
The equations are recursions, we update when new $y_t$ is observed.

# Signal Extraction for Nile Data: filtered estimate of level



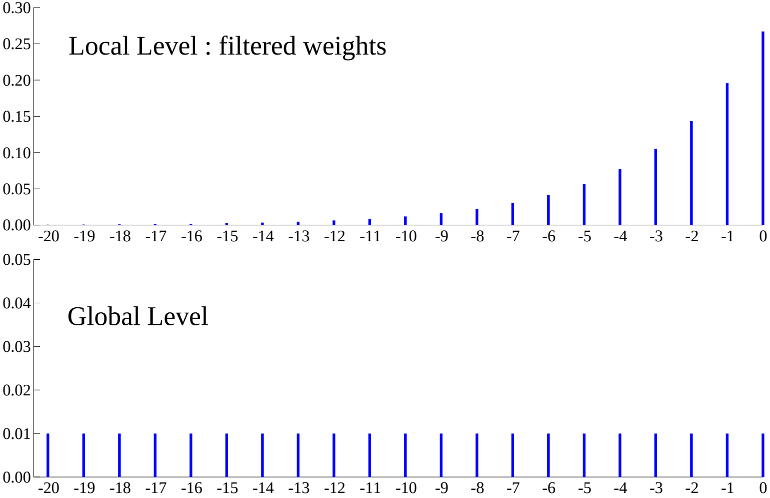
# Observation weights

We show next that $a_{t|t}$ is a weighted sum of past observations :

$$\begin{aligned} a_{t|t} &= a_t + k_t v_t \;=\; a_t + k_t(y_t - a_t) \\ &= k_t y_t + (1-k_t)a_t \\ &= k_t y_t + (1-k_t)a_{t-1} + (1-k_t)k_{t-1}(y_{t-1} - a_{t-1}) \\ &= k_t y_t + k_{t-1}(1-k_t)y_{t-1} + (1-k_t)(1-k_{t-1})a_{t-1} \\ &\vdots \\ &= k_t y_t + \sum_{j=1}^{t-1} w_{t,j} y_{t-j}, \qquad w_{t,j} = k_{t-j} \prod_{m=0}^{j-1}(1-k_{t-m}). \end{aligned}$$

Since $0 < k_t < 1$, the weights are decaying in $j$.
A larger $j$ implies that $y_{t-j}$ is more remote from $t$.

# Signal Extraction for Nile Data: observation weights



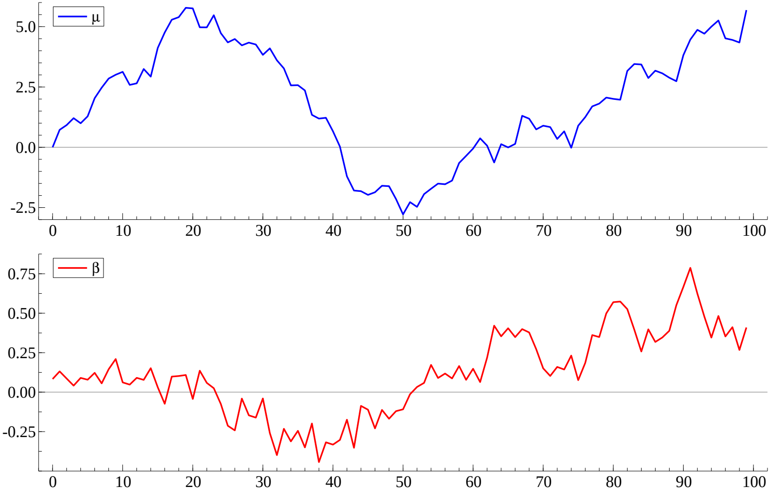
# Local Linear Trend Model

The LLT model extends the LL model with a slope:

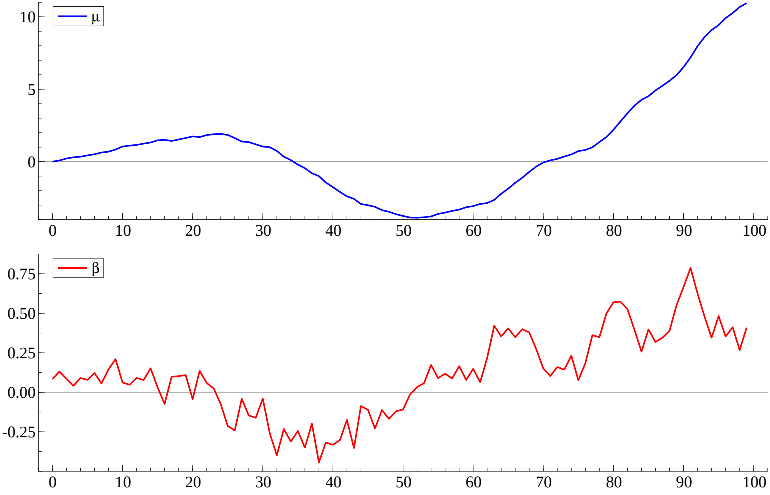
$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{NID}(0, \sigma^2_\varepsilon), \\ \mu_{t+1} &= \beta_t + \mu_t + \eta_t, & \eta_t &\sim \mathcal{NID}(0, \sigma^2_\eta), \\ \beta_{t+1} &= \beta_t + \xi_t, & \xi_t &\sim \mathcal{NID}(0, \sigma^2_\xi). \end{aligned}$$

- All disturbances are independent at all lags and leads;
- Initial distributions $\beta_1, \mu_1$ need to specified;
- If $\sigma^2_\xi = 0$ the trend is a random walk with constant drift $\beta_1$; (For $\beta_1 = 0$ the model reduces to a LL model.)
- If additionally $\sigma^2_\eta = 0$ the trend is a straight line with slope $\beta_1$ and intercept $\mu_1$;
- If $\sigma^2_\xi > 0$ but $\sigma^2_\eta = 0$, the trend is a smooth curve, or an Integrated Random Walk;

# Trend and Slope in LLT Model



# Trend and Slope in Integrated Random Walk Model



# Local Linear Trend Model

- The LLT model can be represented as the ARIMA(0,2,2) model, please verify this;
- The estimation methodology is the same as for the LL model;
- It requires the general state space methods;
- LLT provides a model for Holt-Winters forecasting;
- The smooth trend model is with $\sigma_\zeta^2 = 0$;
- Smoother trend models can be obtained by higher-order Random Walk processes:

$$\Delta^d \mu_t = \eta_t$$

and with $y_t = \mu_t + \varepsilon_t$.

# Seasonal Effects

We have seen specifications for $\mu_t$ in the basic model

$$y_t = \mu_t + \gamma_t + \varepsilon_t.$$

Now we will consider the seasonal term $\gamma_t$. Let $s$ denote the number of 'seasons' in the data:

- $s = 12$ for monthly data,
- $s = 4$ for quarterly data,
- $s = 7$ for daily data when modelling a weekly pattern.

# Dummy Seasonal

The simplest way to model seasonal effects is by using dummy variables. The effect summed over the seasons should equal zero:

$$\gamma_{t+1} = -\sum_{j=1}^{s-1} \gamma_{t+1-j}.$$

To allow the pattern to change over time, we introduce a new disturbance term:

$$\gamma_{t+1} = -\sum_{j=1}^{s-1} \gamma_{t+1-j} + \omega_t, \qquad \omega_t \sim \mathcal{NID}(0, \sigma_\omega^2).$$

The expectation of the sum of the seasonal effects is zero.

## Trigonometric Seasonal

Defining $\gamma_{jt}$ as the effect of season $j$ at time $t$, an alternative specification for the seasonal pattern is

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{jt},$$

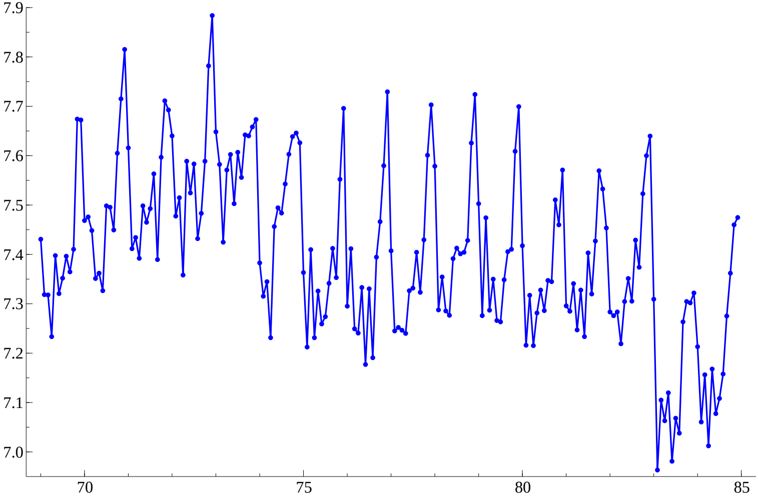
$$\gamma_{j,t+1} = \gamma_{jt} \cos \lambda_j + \gamma_{jt}^* \sin \lambda_j + \omega_{jt},$$

$$\gamma_{j,t+1}^* = -\gamma_{jt} \sin \lambda_j + \gamma_{jt}^* \cos \lambda_j + \omega_{jt}^*,$$

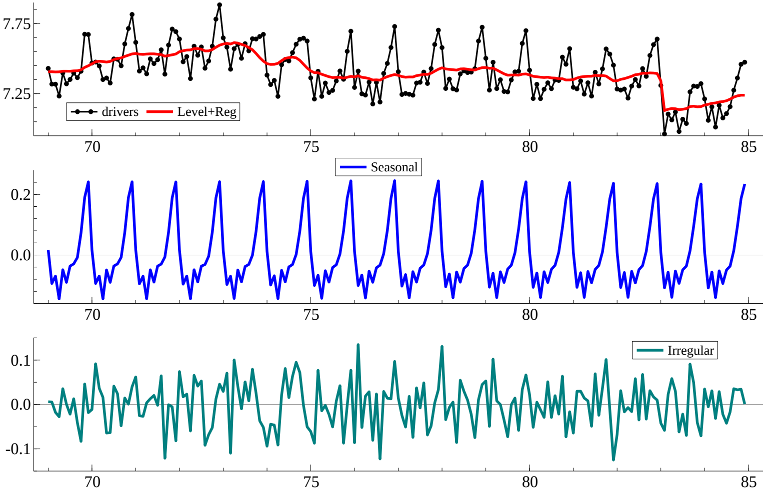
$$\omega_{jt}, \omega_{jt}^* \sim \mathcal{NID}(0, \sigma_\omega^2), \qquad \lambda_j = 2\pi j/s.$$

- Without the disturbance, the trigonometric specification is identical to the deterministic dummy specification.
- The autocorrelation in the trigonometric specification lasts through more lags: changes occur in a smoother way;

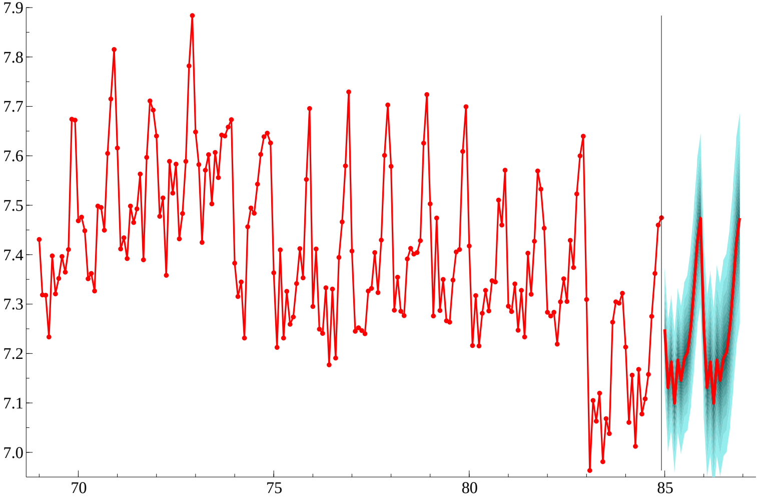
# Seatbelt Law



# Seatbelt Law: decomposition



# Seatbelt Law: forecasting



# Textbooks


- A.C.Harvey (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press
- G.Kitagawa & W.Gersch (1996). *Smoothness Priors Analysis of Time Series*. Springer-Verlag
- J.Harrison & M.West (1997). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag
- J.Durbin & S.J.Koopman (2012). *Time Series Analysis by State Space Methods, Second Edition*. Oxford University Press
- J.J.F.Commandeur & S.J.Koopman (2007). *An Introduction to State Space Time Series Analysis*. Oxford University Press

# Exercises

1 Consider the Local Level model (see slides, see DK chapter 2).

- Reduced form is ARIMA(0,1,1) process. Derive the relationship between signal-to-noise ratio $q$ of LL model and the $\theta$ coefficient of the ARIMA model;
- Derive the reduced form in the case $\eta_t = \sqrt{q}\varepsilon_t$ and notice the difference in the general case.
- Give the elements of the mean vector and variance matrix of $y = (y_1, \ldots, y_n)'$ when $y_t$ is generated by a LL model for $t = 1, \ldots, n$.

# Exercises

2 Consider the stationary time series model

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$
$$\mu_{t+1} = \phi\mu_t + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2),$$

with autoregressive coefficient $|\phi| < 1$ and variances $\sigma_\varepsilon^2 > 0$ and $\sigma_\eta^2 > 0$. The disturbances $\varepsilon_t$ and $\eta_s$ are independent of each other for all $t, s = 1, \ldots, n$.

- Explore the dynamic properties of $y_t$ (mean, variance, autocovariances, autocorrelations).
- Assume that parameters $\phi$, $\sigma_\varepsilon^2$ and $\sigma_\eta^2$ are given. Develop the Kalman filter recursions for this model.
- Propose initial values for mean and variance of the autoregressive component $\mu_t$, that is, $\mu_1 \sim \mathcal{N}(a_1, p_1)$ and propose values for $a_1$ and $p_1$.

# Assignment

3 Consider the Local Level model (see slides, see DK chapter 2).

- Implement the Kalman filter for the Local Level model in a computer program.
- Apply the Kalman filter to the Nile data (Nile data is part of DK book data, see page 11 of these slides)
- Replicate the Figure on page 34 of these slides.

# Appendix – Taylor series

The Taylor expansion for function $f(x)$ around some value $x^*$ is

$$f(x) = f(x = x^*) + f'(x = x^*)[x - x^*] + \frac{1}{2} f''(x = x^*)[x - x^*]^2 + \ldots,$$

where

$$f'(x) = \frac{\partial f(x)}{\partial x}, \qquad f''(x) = \frac{\partial^2 f(x)}{\partial x \partial x},$$

and $g(x = x^*)$ means that we evaluate function $g(x)$ at $x = x^*$.

*Example*: consider $f(x) = \log(1 + x)$ with $f'(x) = (1 + x)^{-1}$ and $f''(x) = -(1 + x)^{-2}$; the expansion of $f(x)$ around $x^* = 0$ is

$$\log(1 + x) = 0 + 1 \cdot (x - 0) + \frac{1}{2}(-1) \cdot (x - 0)^2 + \ldots = x - \frac{1}{2}x^2 + \ldots$$

Notice that $f(x = 0) = 0$, $f'(x = 0) = 1$ and $f''(x = 0) = -1$. For small enough $x$ (when $x$ is close to $x^* = 0$), we have

$$\log(1 + x) \approx x.$$

Check: $\log(1.01) = .00995 \approx 0.01$ and $\log(1.1) = 0.0953 \approx 0.1$.

# Appendix – Percentage growth

Observation at time $t$ is $y_t$ and observation at time $t-1$ is $y_{t-1}$.

We define rate $r_t$ as the proportional change of $y_t$ wrt $y_{t-1}$, that is

$$r_t = \frac{y_t - y_{t-1}}{y_{t-1}} \Rightarrow y_t - y_{t-1} = y_{t-1} \cdot r_t \Rightarrow y_t = y_{t-1} \cdot (1 + r_t).$$

We notice that $r_t$ can be positive and negative !

When we take logs of $y_t = y_{t-1} \cdot (1 + r_t)$, we obtain

$$\log y_t = \log y_{t-1} + \log(1 + r_t) \Rightarrow \log y_t - \log y_{t-1} = \log(1 + r_t) \Rightarrow$$

$$\Delta \log y_t = \log(1 + r_t).$$

Since $\log(1 + r_t) \approx r_t$, see previous slide, when $r_t$ is small, we have

$$r_t \approx \Delta \log y_t.$$

The percentage growth is defined as $100 \times r_t \approx 100 \cdot \Delta \log y_t$.

# Appendix – Lag operators and polynomials

- Lag operator $Ly_t = y_{t-1}$, $L^\tau y_t = y_{t-\tau}$, for $\tau = 1, 2, 3, \ldots$
- Difference operator $\Delta y_t = (1 - L)y_t = y_t - y_{t-1}$
- Autoregressive polynomial $\phi(L)y_t = (1 - \phi L)y_t = y_t - \phi y_{t-1}$
- Other polynomial $\theta(L)\varepsilon_t = (1 + \theta L)\varepsilon_t = \varepsilon_t + \theta\varepsilon_{t-1}$
- Second difference
  $\Delta^2 y_t = \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2}$
- Seasonal difference $\Delta_s y_t = y_t - y_{t-s}$ for typical
  $s = 2, 4, 7, 12, 52$
- Seasonal sum operator
  $S(L)y_t = (1 + L + L^2 + \ldots + L^{s-1})y_t = y_t + y_{t-1} + \ldots + y_{t-s+1}$
- Show that $\Delta\, S(L) = \Delta_s$.